Volume 99, Number 4, July–August 1994
Journal of Research of the National Institute of Standards and Technology

[J. Res. Natl. Inst. Stand. Technol. **99**, 539 (1994)]

On the Convergence of the Number of Exceedances of Nonstationary Normal Sequences

Volume 99 Number 4 July–August 1994

J. Hüsler

University of Bern,
Bern, Switzerland

and

M. Kratz

University Paris VI,
Paris, France

It is known that the number of exceedances of normal sequences is asymptotically a Poisson random variable, under certain restrictions. We analyze the rate of convergence to the Poisson limit and extend the result known in the stationary case to nonstationary normal sequences by using the Stein-Chen method. In addition, we consider the cases of exceedances of a constant level as well as of a particular nonconstant level.

Key words: exceedances; nonstationary; normal sequences; rate of convergence; Stein-Chen method.

Accepted: March 22, 1994

1. Introduction and Result

The extreme value theory of Gaussian sequences has interested many authors, for instance Refs. [1,4,7,8], dealing with the limit distribution of the suitably normalized extreme value.

Let $\{X_i, i \geq 1\}$ be a standardized normal sequence with correlations $E(X_iX_j) = r_{ij}$, $i,j \geq 1$, and $\Phi(\cdot)$ the distribution function of X_i.

Let

$$N_n = \sum_{i=1}^{n} 1(X_i > u_{ni})$$

denote the number of exceedances of a boundary given by a triangular array $\{u_{ni}, i \leq n, n \geq 1\}$. Then it was also found that N_n converges in distribution to a random variable having a Poisson distribution $\mathbb{P}(\lambda_n)$, if the mean number of exceedances $\lambda_n = \sum_{i \leq n}(1 - \Phi(u_{ni}))$ remains bounded (cf. Ref. [4]).

For practical use of the asymptotic theory, it is rather important to know the rate of convergence or at least some upper bound for this rate.

For the stationary case, results on the rate of convergence have been obtained for instance by Refs. [2,3,9,10]. The aim of this paper is to give an upper bound for the total variation distance d_{tv} between N_n and $\mathbb{P}(\lambda_n)$, in the nonstationary case, extending the results of these mentioned papers.

Suppose that for some sequence ρ_n: $|r_{ij}| \leq \rho_{|i-j|}$ for $i \neq j$, and that the two conditions

$$\rho_n < 1 \text{ for all } n \geq 1 \tag{1}$$

$$\rho_k \leq A/\log k, k \geq 2, \text{ for some constant } A \tag{2}$$

are satisfied. Define ρ as $\rho = \max(0, r_{ij}, i \neq j) < 1$.

In addition, we assume that the boundary values tend uniformly to ∞:

$$u_{n,\min} = \min_{1 \leq i \leq n} u_{ni} \to \infty \text{ as } n \to \infty. \tag{3}$$

The exceedances of a constant boundary $u_{ni}=u_n$, $1 \leq i \leq n$, are considered first, where only the tools given in Ref. [3] are used. We show in the second result that the method of Ref. [3] can be used also for nonconstant boundaries $\{u_{ni}\}$. But these boundaries are restricted such that the condition

$$\limsup_{n\to\infty} n(1-\Phi(u_{n,\min}))<C<\infty \tag{4}$$

holds. If we want to extend the results to a more general class of boundaries such that only

$$\limsup_{n\to\infty} \lambda_n<C<\infty \tag{5}$$

holds, we need to combine the method developed by Ref. [4] with that of Ref. [3] to get satisfactory results (see Ref. [6]).

Our first result for boundaries which are constant for fixed n, shows that the given upper bound of the rate of convergence depends mainly on the largest positive correlation value ρ.

Theorem 1: *Let $\{X_i, i \geq 1\}$ be a standardized nonstationary normal sequence with correlations $\{r_{ij}, i,j \geq 1\}$. Suppose that $|r_{ij}| \leq \rho_{|i-j|}$ for $i \neq j$, such that Eqs. (1) and (2) hold. Let the boundary values $\{u_{ni}=u_n, 1 \leq i \leq n\}$ and λ_n be real values with $\lambda_n = n(1-\Phi(u_n))$. Suppose that $\lambda_n \leq C < \infty$, for some constant C. Then as $n \to \infty$*

$$d_{tv}(N_n, \mathbb{P}(\lambda_n)) = O\Big(n^{-\frac{1-\rho}{1+\rho}} \cdot (\log n)^{-\frac{\rho}{1+\rho}}$$

$$+\frac{\log n}{n^2}\sum_{k=1}^{n-1}(n-k)\rho_k\Big). \tag{6}$$

This extends the result of Ref. [3] showing that for a constant boundary their upper bound of the rate of convergence in the stationary case holds also in the nonstationary case.

Theorem 2: *Let $\{X_i, i \geq 1\}$ be a standardized nonstationary normal sequence with correlations $\{r_{ij}, i,j \geq 1\}$ as in Theorem 1 satisfying Eqs. (1) and (2). Suppose that the boundary values $\{u_{ni}, 1 \leq i \leq n, n \geq 1\}$ are such that Eq. (4) holds. Then Eq. (6) holds.*

The first term of Eq. (6) dominates the rate of convergence in cases with $\Sigma_{k\geq 1}\rho_k < \infty$ and $\rho > 0$.

Then the rate of convergence depends only on the lowest value $u_{n,\min}$ of the particular boundary u_{ni} and also on the largest positive correlation ρ. It extends naturally the results of the stationary case with boundary values which are constant for fixed n. This rate is only good if $u_{n,\min}$ is not a uniquely low value, which is supposed for reasonable boundaries. For the case $u_{n,\min}$ is uniquely low, the rate can be improved.

2. Proof

The proof of Theorem 1 is an adaption of that used by Ref. [3] in the stationary case. We use the following lemma which is a straightforward extended version of Lemma 3.4 of Ref. [3].

Lemma 1: *Suppose that*

$$(X_i, X_j) \stackrel{d}{=} N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1 & r_{ij}\\ r_{ij} & 1\end{pmatrix}\right).$$

Define $Z_i = 1(X_i > u_{ni})$ for some boundary values u_{ni} where

$$1-\Phi(u_{ni}) \leq C/n$$

for some finite constant C.
Then for some constant K depending on C only and for all $n \geq 2$:

i) If $0 \leq r_{ij} < 1$, then

$$0 \leq cov(Z_i, Z_j) \leq K \cdot \frac{1}{\sqrt{1-r_{ij}}} n^{-2+\frac{2r_{ij}}{1+r_{ij}}} (\log n)^{-\frac{r_{ij}}{1+r_{ij}}}$$

$$\leq K n^{-2}\left(\frac{n^2}{\log n}\right)^{\frac{\rho_{|i-j|}}{1+\rho_{|i-j|}}}$$

ii) If $0 \leq r_{ij} \leq 1$, then

$$0 \leq cov(Z_i, Z_j) \leq K\frac{r_{ij}\log n}{n^2} e^{2r_{ij}\log n}$$

$$\leq K\frac{\rho_{|i-j|}\log n}{n^2} e^{2\rho_{|i-j|}\log n}$$

iii) If $-1 < r_{ij} \leq 0$, then

$$0 \geq cov(Z_i, Z_j) \geq -K\frac{|r_{ij}|\log n}{n^2} \geq -K\frac{\rho_{|i-j|}\log n}{n^2}$$

iv) If $-1 \leq r_{ij} \leq 0$, then

$$0 \geq cov(Z_i, Z_j) \geq -K\frac{1}{n^2}$$

We need for Theorem 2 an extension of Lemma 1.

Lemma 2: *Suppose that* (X_i, X_j) *and* Z_i *are as in Lemma 1. For any* i,j, *define* $u_{nij} = \min(u_{ni}, u_{nj})$ *and* $v_{nij} = \max(u_{ni}, u_{nj})$.

Then for some constant K depending only on C and for all $n \geq 2$:

i) If $0 \leq r_{ij} \leq 1$, *then*

$$0 \leq cov(Z_i, Z_j) \leq K \frac{1}{\sqrt{1-r_{ij}}} \left(\frac{\varphi(u_{nij})}{u_{nij}}\right)^{\frac{2}{1+r_{ij}}} \cdot u_{nij}^{-\frac{2r_{ij}}{1+r_{ij}}}$$

with $K \geq (2\pi)^{-\frac{r_{ij}}{1+r_{ij}}} \cdot (1+r_{ij})^{3/2}$.

ii) If $0 \leq r_{ij} \leq 1$, *then*

$$0 \leq cov(Z_i, Z_j) \leq K\, r_{ij}\, \varphi(u_{nij}) \Big(\varphi(v_{nij})\Big)^{\frac{1-r_{ij}}{1+r_{ij}}}$$

iii) If $-1 \leq r_{ij} \leq 0$, *then*

$$0 \geq cov(Z_i, Z_j) \geq -(1-\Phi(u_{nij}))(1-\Phi(v_{nij}))$$

$$\geq -(1-\Phi(u_{nij}))^2$$

iv) If $-1 \leq r_{ij} \leq 0$, *then*

$$0 \leq |cov(Z_i, Z_j)| \leq K \min(1, |r_{ij}| u_{nij} v_{nij} + r_{ij}^2 v_{nij}^2)$$

$$(1-\Phi(u_{nij})) \cdot (1-\Phi(v_{nij}))$$

(The proof of this lemma is given in [6]).

Theorem 1 follows also by Theorem 2. Therefore, we prove now Theorem 2 by using the method of Ref. [3].

By Theorem 3.1 of Ref. [3], we have

$$d_{tv}(N_n, I\!P(\lambda_n)) \leq \frac{1-e^{-\lambda_n}}{\lambda_n} \left(\frac{\lambda_n^2}{n} + \sum_{i \neq j} |cov(Z_i, Z_j)|\right). \quad (7)$$

If $\rho = 0$, i.e. if $r_{ij} \leq 0$ for $i \neq j$, then using Lemma 2(iv) for the second term of Eq. (7), we get the second term of Eq. (6) which dominates the first one in this case, if $\rho_k > 0$ for some k. Obviously, if $\rho_k = 0$ for all k, then the result holds, since the second term in Eq. (7) is 0.

Thus suppose from now on that $\rho > 0$.

Because of Eq. (2) the sum

$$S_n = \sum_{1 \leq i < j \leq n} |cov(Z_i, Z_j)| = \sum_{1 \leq i < j \leq n} c_{ij}$$

is split up into three parts, by using $\delta > 0$ such that

$$3\delta < \frac{\rho}{1+\rho}$$

There are only finitely many k's with $\rho_k > \delta$. This will be treated first as Case (i). For indices k with $\rho_k \leq \delta$, we distinguish Case (ii) with $k < n^\delta$ and Case (iii) with $k \geq n^\delta$.

Case (i): Each term c_{ij} of the sum S_n is bounded above by

$$K n^{-2/(1+\rho)} (\log n)^{-\rho/(1+\rho)}$$

if $r_{ij} \geq 0$ by Lemma 2(i)
or bounded by

$$K n^{-2}$$

if $r_{ij} < 0$ by Lemma 2(iii).
Since there are finitely many k's with $\rho_k > \delta$, the number of terms c_{ij} with $|i-j| = k$ and $\rho_k > \delta$ is of the order $O(n)$. Hence the sum on these terms is bounded by

$$K n^{-(1-\rho)/(1+\rho)} (\log n)^{-\rho/(1+\rho)} + K n^{-1}.$$

Case (ii): There are at most n^δ terms ρ_k such that $0 \leq \rho_k \leq \delta$ and $k < n^\delta$. Hence there are at most $O(n^{1+\delta})$ terms c_{ij} with such a $k = |i-j|$. But each such term c_{ij} of S_n is bounded by

$$K n^{-2/(1+\delta)} (\log n)^{-\delta/(1+\delta)} \leq K n^{-2+2\delta}$$

if $r_{ij} \geq 0$ by Lemma 2(i)
or bounded by

$$K n^{-2}$$

if $r_{ij} < 0$ by Lemma 2(iv).
Then the sum of these terms c_{ij} is bounded by $O(n^{-1+3\delta})$.

Case (iii): Finally we consider the terms such that $\rho_k \leq \delta$, with $k \geq n^\delta$. Note that we have

$$0 \leq \rho_k \leq \frac{A}{\log k} \leq \frac{A}{\delta \log n}. \quad (8)$$

Each such term c_{ij} of S_n gives a contribution

$$K\,\rho_k\left(\frac{\sqrt{\log n}}{n}\right)^{1+\frac{1-\rho_k}{1+\rho_k}} \leqslant K\,\rho_k\frac{\log n}{n^2}$$

if $r_{ij} \geqslant 0$ by Lemma 2(ii) and by using Eq. (8)

$$K\,\rho_k\left(\frac{\log n}{n^2}\right)$$

if $r_{ij} < 0$ by Lemma 2(iv).
Taking now the sum on all terms (i,j) with $|i-j| \geqslant n^{\delta}$, we get the second term of Eq. (6).

Finally, adding up all these upper bounds of Cases (i), (ii), and (iii), the result Eq. (6) of Theorem 2 follows.

3. References

[1] S. M. Berman, Limit theorems for the maximum term in stationary sequences, Ann. Math. Statist. **33**, 502–516 (1964).
[2] P. Hall, The rate of convergence of normal extremes, J. Appl. Probab. **16**, 433–439 (1979).
[3] L. Holst and S. Janson, Poisson approximation using the Stein-Chen method and coupling: number of exceedances of Gaussian random variables, Ann. Probab. **18**, 713–723 (1990).
[4] J. Hüsler, Asymptotic approximations of crossing probabilities of random sequences, Z. Wahrschcinlichkeitstheorie verw. Geb. **63**, 257–270 (1983).
[5] J. Hüsler, Extreme values of non-stationary random sequences, J. Appl. Probab. **23**, 937–950 (1986).
[6] J. Hüsler and M. F. Kratz, Rate of convergence of the number of exceedances of nonstationary normal sequences. Preprint (1993).
[7] M. R. Leadbetter, G. Lindgren, and H. Rootzén, Extremes and related properties of random sequences and processes, Springer Series in Statistics, Springer, New York (1983).
[8] V. I. Piterbarg, Asymptotic expansions for the probability of large excursions of Gaussian processes, Soviet Math. Dokl. **19**, 1279–1283 (1978).
[9] H. Rootzén, The rate of convergence of extremes of stationary random sequences, Adv. Appl. Probab. **15**, 54–80 (1983).
[10] R. Smith, Extreme value theory for dependent sequences via the Stein-Chen method of Poisson approximation, Stoch. Process. Appl. **30**, 317–327 (1988).

About the authors: *Jürg Hüsler is Professor at the Department of Mathematical Statistics of the University of Bern. Marie Kratz is now Maître de Conférences at the Department of Mathematics and Computer Science of the University René Descartes at Paris.*